RSAC 2025 Conference

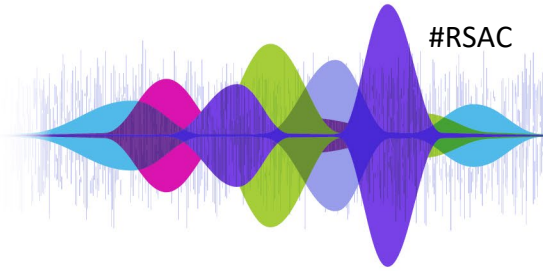
Many Voices.
One Community.

SESSION ID: PART3-W01

Fast-Track Security Automation with LLM-Driven Workflows

Steve Povolny

Senior Director, Security Research
https://www.linkedin.com/in/steve-povolny-595a776



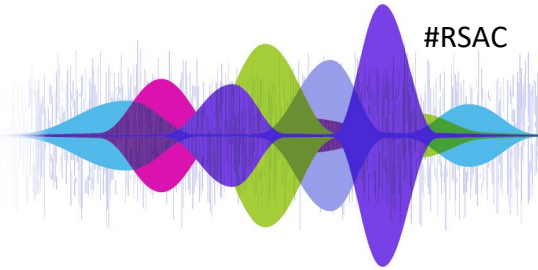
Disclaimer

Presentations are intended for educational purposes only and do not replace independent professional judgment. Statements of fact and opinions expressed are those of the presenters individually and, unless expressly stated to the contrary, are not the opinion or position of RSA Conference LLC or any other co-sponsors. RSA Conference LLC does not endorse or approve, and assumes no responsibility for, the content, accuracy or completeness of the information presented.

Attendees should note that sessions may be audio- or video-recorded and may be published in various media, including print, audio and video formats without further notice. The presentation template and any media capture are subject to copyright protection.

© 2025 RSA Conference LLC or its affiliates. The RSAC and RSAC CONFERENCE logos and other trademarks are proprietary. All rights reserved.

By the end of this session, you will learn:



The top tools and automation frameworks available today

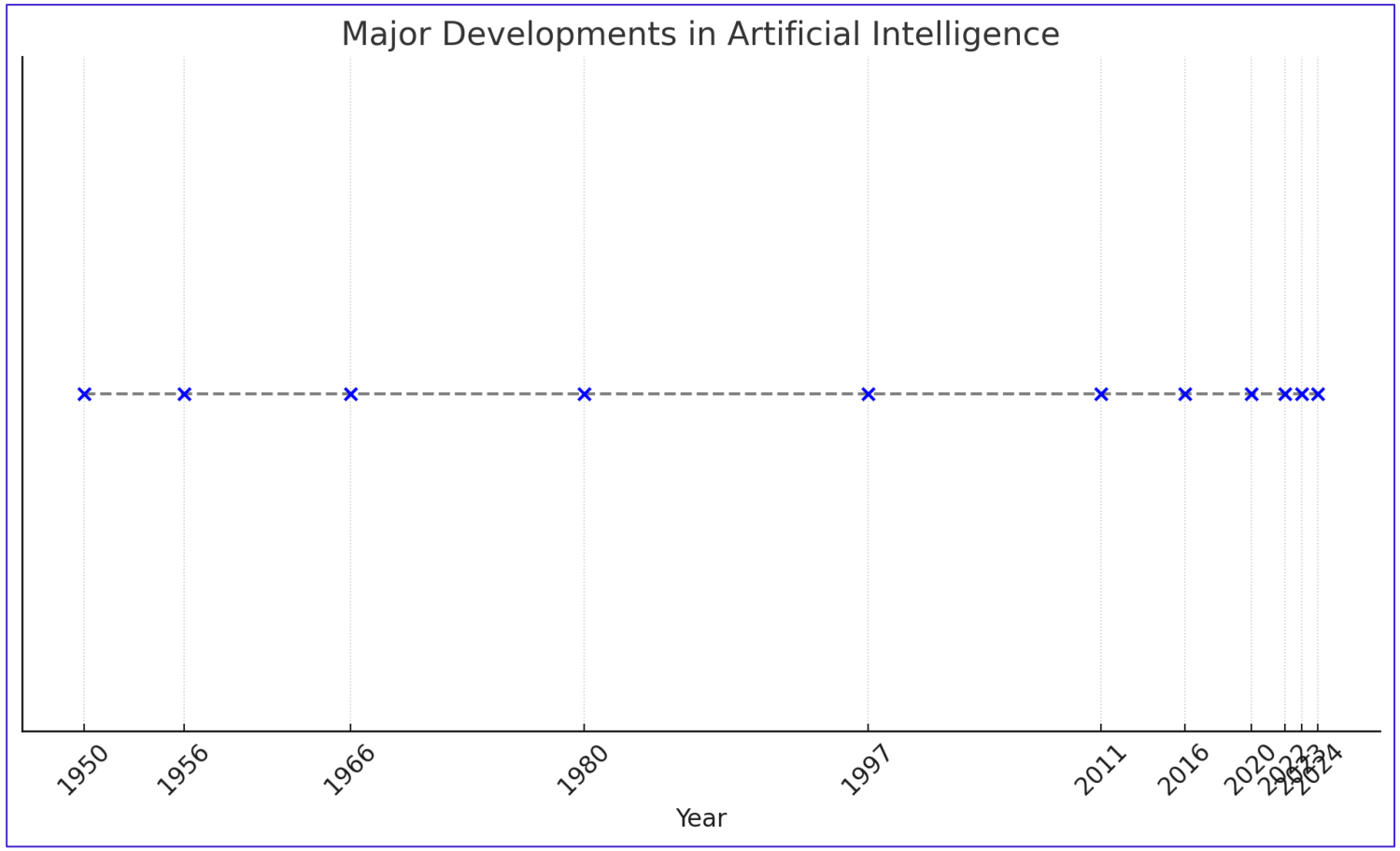
Real-world applications of LLMs in security operations

Best practices for prompt engineering

How to integrate AI and LLMs into security automation

Lessons learned, challenges, and practical insights

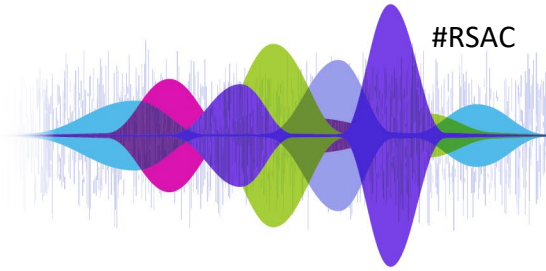
AI’s Exponential Growth



Major AI Developments

- **1950**: Alan Turing's *Computing Machinery and Intelligence*
- **1956**: Dartmouth Conference: AI as a field is born
- **1966**: ELIZA, the first chatbot, is created
- **1980**: Expert Systems gain commercial success
- **1997**: IBM's Deep Blue defeats chess champion Garry Kasparov
- **2011**: IBM Watson wins *Jeopardy!*
- **2016**: AlphaGo defeats world Go champion Lee Sedol
- **2020**: GPT-3 is released by OpenAI
- **2022**: ChatGPT (GPT-3.5) released, AI adoption skyrockets
- **2023**: Multimodal AI expands capabilities; reasoning models improve
- **2024**: Open-source LLMs proliferate; DeepSeek challenges cost paradigm; AI agents gain autonomy

A Brief History of LLMs



Audience Questions

What year was ChatGPT first released?

2022

What month?

November

What year was the first iteration of GPT released?

June, 2018
GPT-1 117M

LLM Tools Available Today

ANTHROP\C

cohere

deepseek

Gemini

Meta

IBM

Microsoft

OpenAI

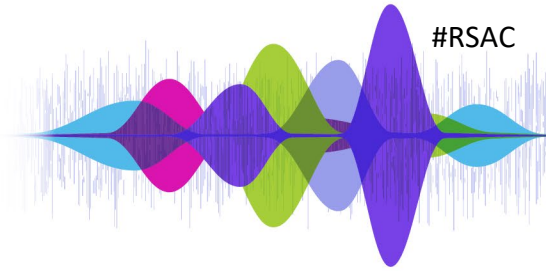
RSAC 2025 Conference

Foundations of LLM Security & Automation

Why Use LLMs for Security?

- ✓ Automate security workflows and responses
- ✓ Reduce analyst workload by automating repetitive tasks
- ✓ Generate contextual insights from security data
- ✓ Improve SOC efficiency with AI-driven recommendations

exabeam



Security Considerations When Using LLMs

Data Privacy & Compliance	→	Avoid exposing sensitive logs or PII
Prompt Injection Risks	→	Prevent adversarial manipulation
Model Bias & Hallucinations	→	Validate LLM responses against ground truth
API Security	→	Secure API integrations with proper authentication

What is Prompt Engineering?

- **The practice of crafting prompts to guide LLM responses**
- **The quality of the prompt directly impacts the accuracy of the response**
- **Examples:**
  - **Basic Prompt:** *“Analyze this security log.”*
  - **Optimized Prompt:** *“Analyze the following log for indicators of compromise (IOCs) related to phishing attempts. Highlight any anomalies.”*

exabeam

Prompt Engineering

- Persona
- Instructions
- Refinements
- Output/Format
- **Controls** (Temperature, etc)
- **Context Data**

A security analyst needs help understanding an incident that was escalated into their case management system. **Your job is to summarize the incident**, explaining what the possible threat is and what actions the analyst may want to take next. **Instead of repeating the details of each event**, explain the connection between them, and how together they may indicate malicious activity.

Please format your summary in the following way:
Overview: Provide a few narrative-driven sentences that tie all the events in the incident together. Be sure to include the event details like ips, hosts, usernames, domains and other relevant fields in the overview. Do not hyperlink urls.

Possible Threats: Provide bullet list explaining what possible threats the incident may indicate.

Next Steps: Provide a bullet list of actions the analyst might consider taking next.

Summarize the following incident:

RAG / Context

- **Retrieval-augmented-generation**
  - “Stuff to make responses better”
- **Structured or unstructured data retrieved from an external source or augmenting the prompt inline**
  - {JSON formatted detections and information pulled from a database or vector store}

{ "data": [{ "approx_log_time": "Aug 16, 2024 12:26 PM", "events": [{ "action": "Allow", "activity": "traffic", "activity_type": "network-traffic", "approxLogTime": 1723811195000000, "bytes_in": 120796645, "bytes_out": 157284696, "customFieldsJSON": "{}", "dest_ip": "44.220.86.240", "dest_port": 60061, "event_code": "3000-0151", "geo_dest_ip": { "city": "ashburn", "country": "us", "isp": "amazon.com inc.", "latitude": "39.0437", "longitude": "-77.4742" },

Tuning Responses for Accuracy and Reliability

Common LLM Issues & Solutions:

- **Hallucinations (Fake Data):**
  - Cross-check responses with external sources
  - Implement confidence scoring mechanisms
- **Excessive verbosity:**
  - Use structured prompts to limit response length
- **Security risks in responses:**
  - Filter and sanitize responses before execution

exabeam

RSAC 2025 Conference

Real-World Applications of LLM-Driven Security Automation

Many Voices.
One Community.

Automation Frameworks for LLM Integration

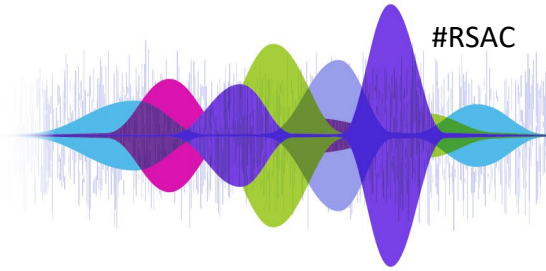
Key Automation Platforms:

- **LangChain**
  Framework for integration with LLMs
- **N8N**
  Secure and AI native automation workflow (nodes built on LC)
- **AutoGen**
  Open-source AI agent building framework
- **Windmill**
  Open-source developer platform and workflow engine
- **Google Vertex Agent Framework**
  Multi-agent development and workflow framework
- **Many more!**

Workflow Automation Concepts

- **Setup** (APIs, Keys, Instances)
- **Nodes** (Trigger, Agent, Code, Data transformation, Flow, Core, HitL)
- **Session pinning, reruns**
- **Prior executions**
- **Execution tracing**
- **Community library of templates**

Real-World SOC Use Case



Agent-driven threat intelligence correlation via chat interface

- A security analyst receives an alert from the SIEM
- LLM automatically enriches the alert with external threat intelligence
- AI summarizes and provides recommended actions
- Analyst verifies and executes the response

RSAC 2025 Conference

Agents and Agentic AI

What's an Agent?

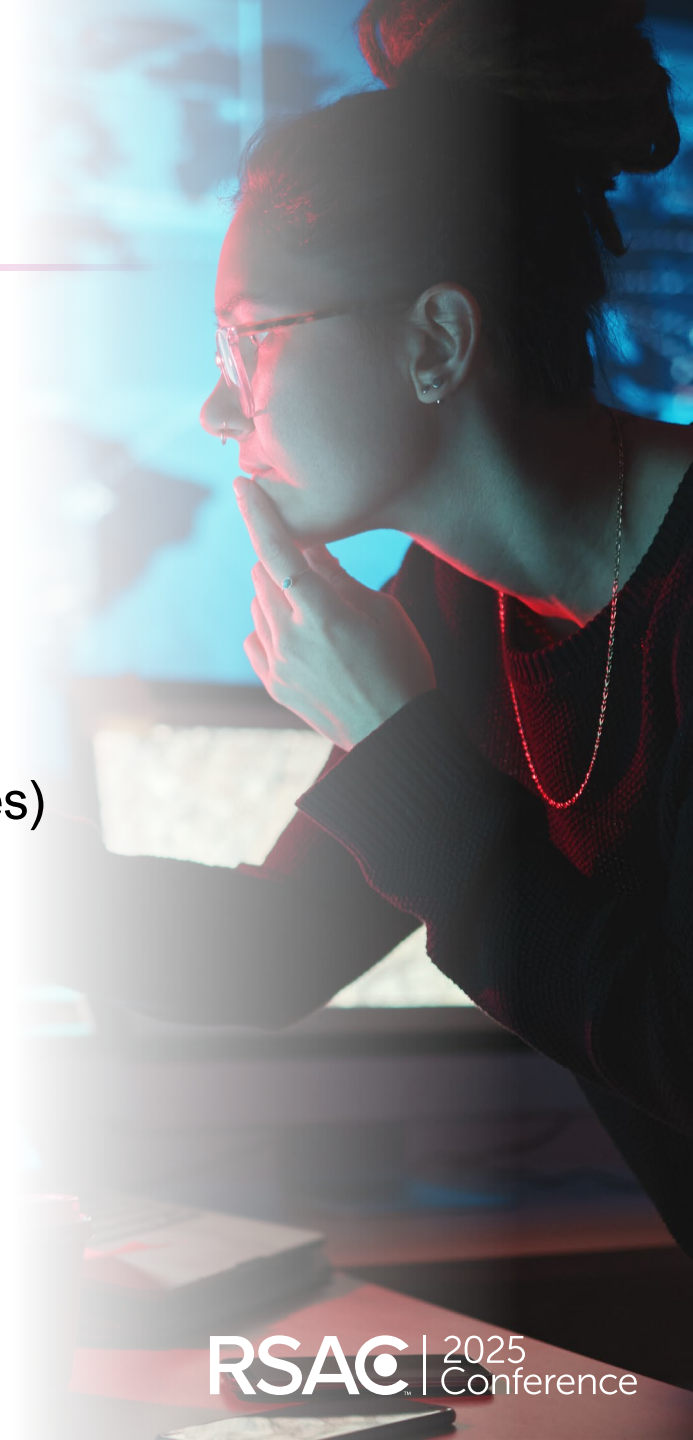
- AI agent is a purpose-driven component that addresses a specific task using a form of AI
- Agentic AI is an autonomous system that can operate independently and can learn and adapt to changing conditions
- Agent: Suggest recipes, ingredients and directions for meal planning
- Agentic AI: Build a weekly meal plan, shop for groceries, schedule pickup, and cook the meal (ok not the last part)

Steve's Chat Agent for Security Operations

- **Workshop Session**
- **Prebuilt flows:**
  LLM agent, Chat agent, Vector Store Write/Read, API request
- **Threat Intelligence Multi-Agent Workflow**
- **Leveraging a SIEM log or detection event**

Configuring Self-Hosted AI Starter Kit

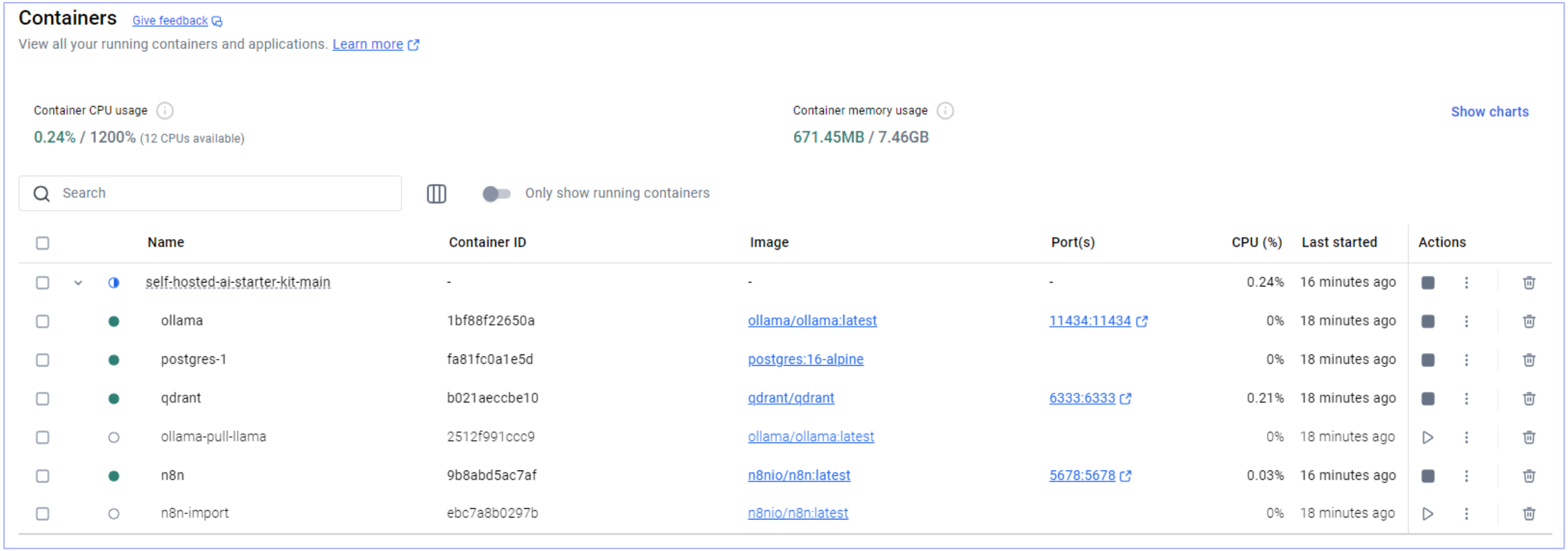
- **Ollama**
  (for open source LLM)
- **N8N**
  (for automation)
- **Docker**
  (for containerization and local instances)
- **Qdrant**
  (for vector store)
- **Postgres**
  (agent memory and workflow storage)



Installation

- **Install Docker**
- **Clone the GitHub repo**
- **Run docker for your setup**
  (AMD, NVIDIA, Non-GPU, etc…)
- **Configure N8N**
  (create a local non-cloud account)

Docker



Containers Give feedback

View all your running containers and applications. Learn more

Container CPU usage
0.24% / 1200% (12 CPUs available)

Container memory usage
671.45MB / 7.46GB

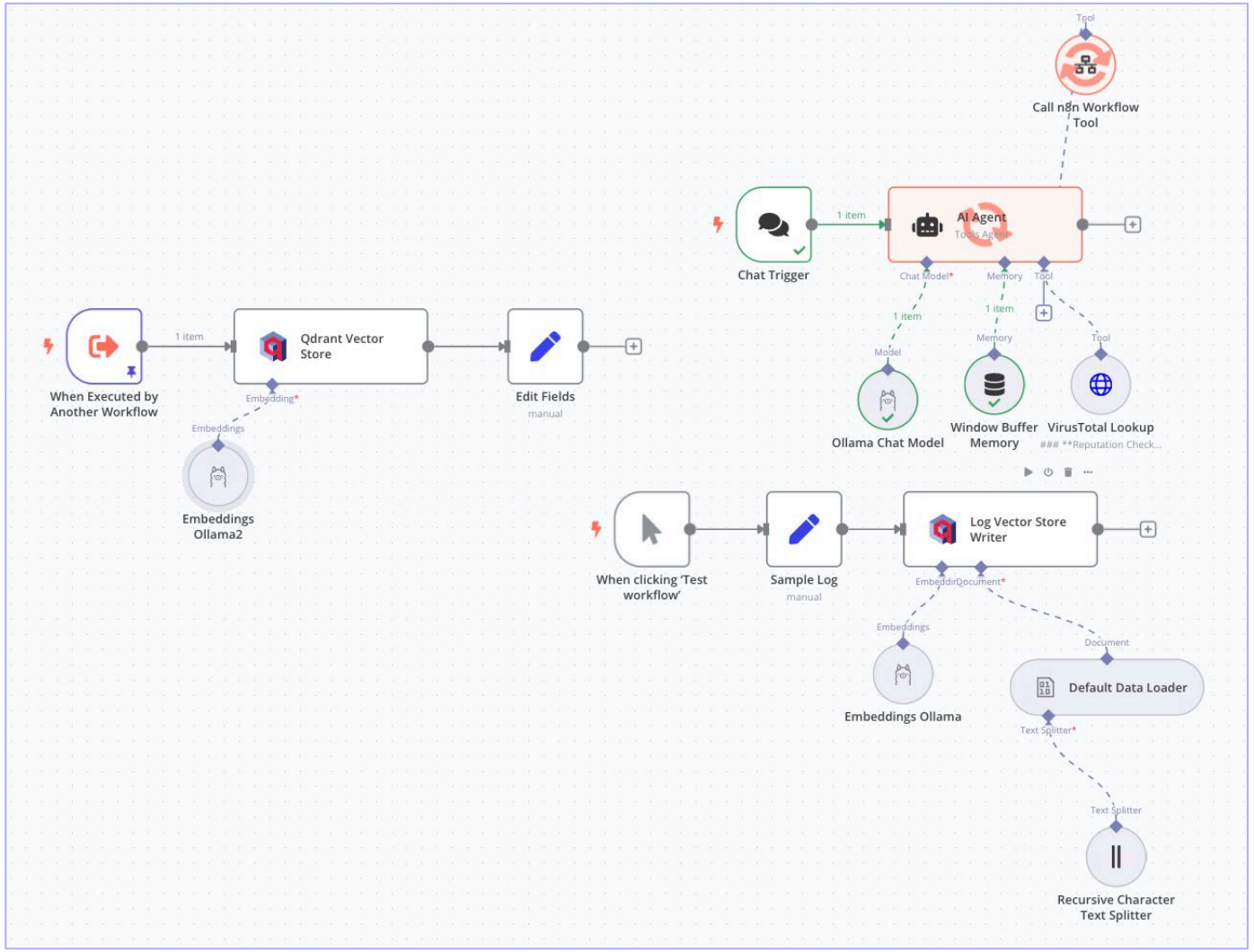
Show charts

Search

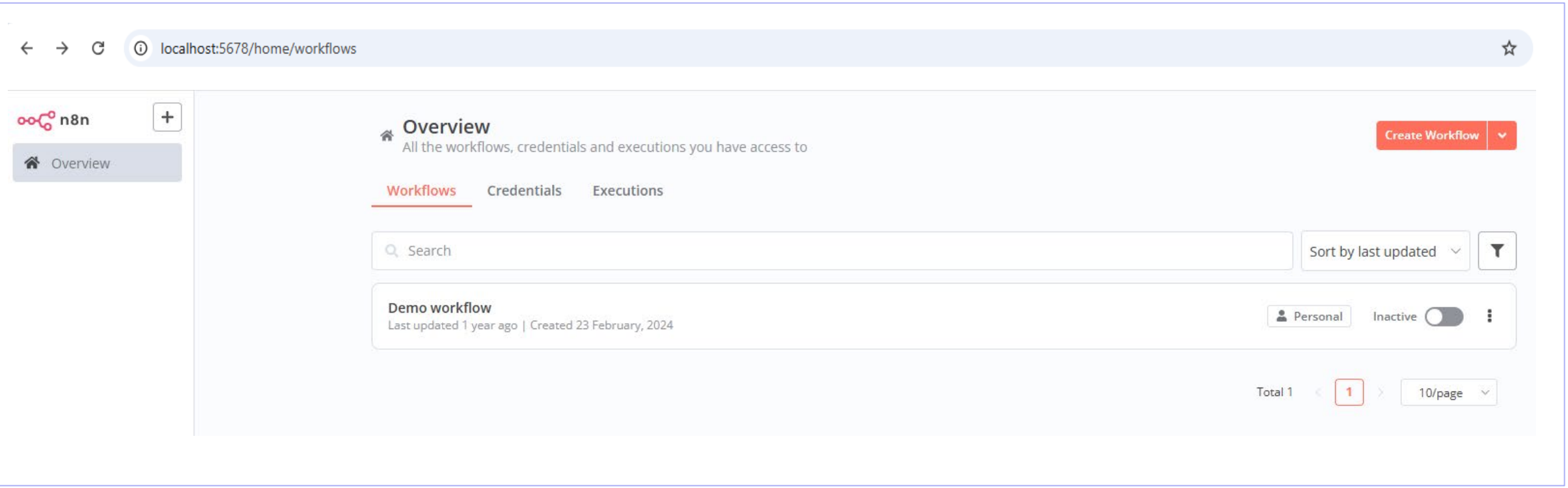
Only show running containers

Name	Container ID	Image	Port(s)	CPU (%)	Last started	Actions
self-hosted-ai-starter-kit-main	-	-	-	0.24%	16 minutes ago	
ollama	1bf88f22650a	ollama/ollama:latest	11434:11434	0%	18 minutes ago	
postgres-1	fa81fc0a1e5d	postgres:16-alpine		0%	18 minutes ago	
qdrant	b021aeccbe10	qdrant/qdrant	6333:6333	0.21%	18 minutes ago	
ollama-pull-llama	2512f991ccc9	ollama/ollama:latest		0%	18 minutes ago	
n8n	9b8abd5ac7af	n8nio/n8n:latest	5678:5678	0.03%	16 minutes ago	
n8n-import	ebc7a8b0297b	n8nio/n8n:latest		0%	18 minutes ago	

Jumping Ahead

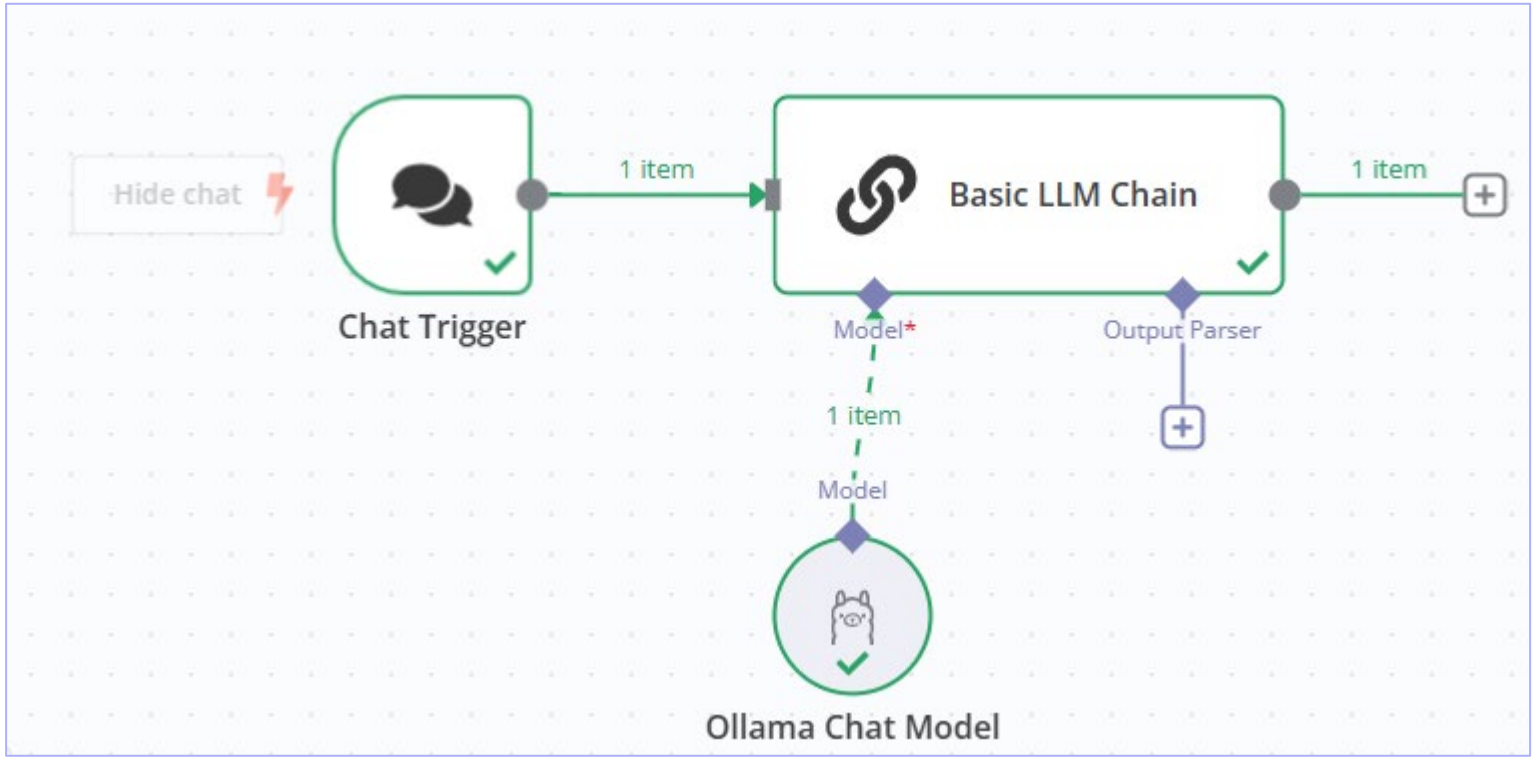


N8N on Localhost:5678

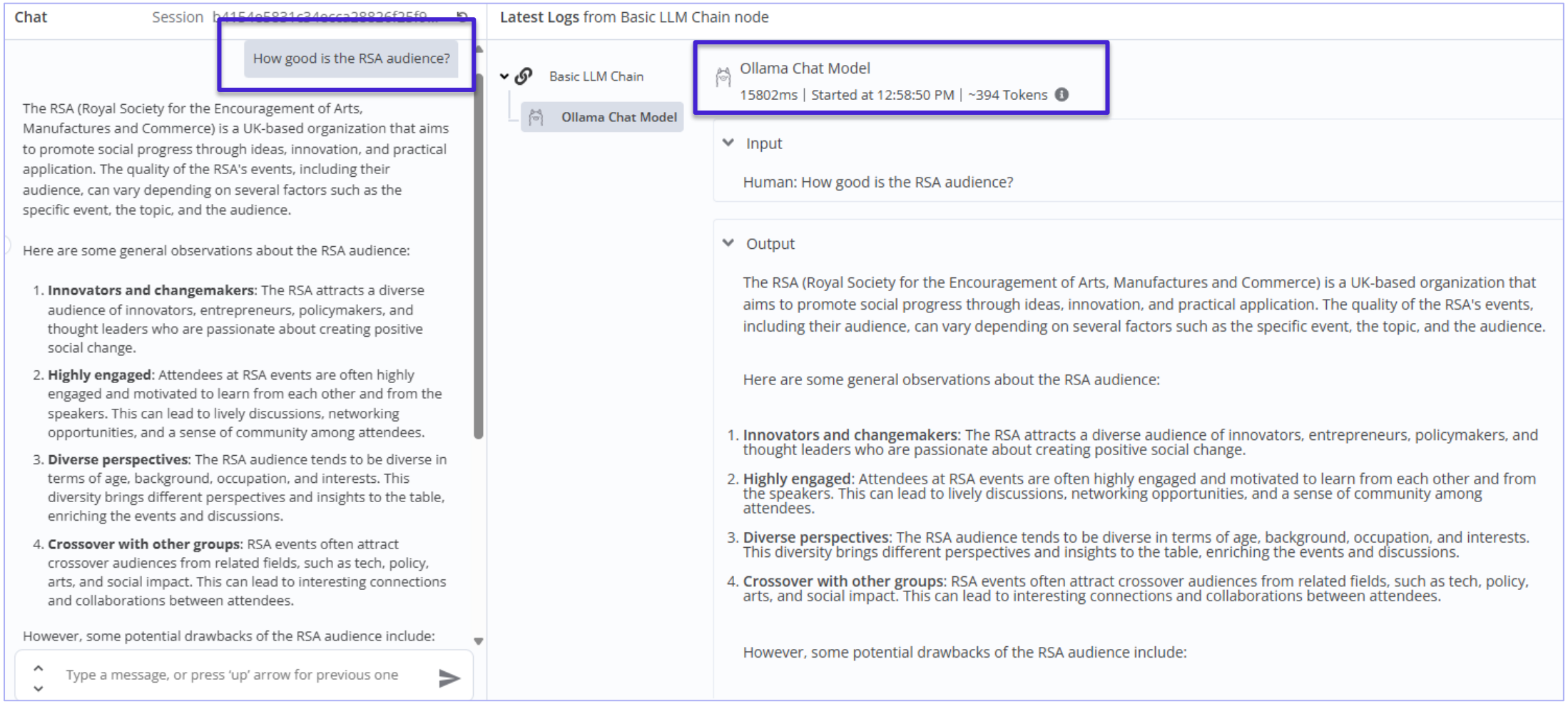


Demo Workflow

Already created with simple chat + ollama LLM integration



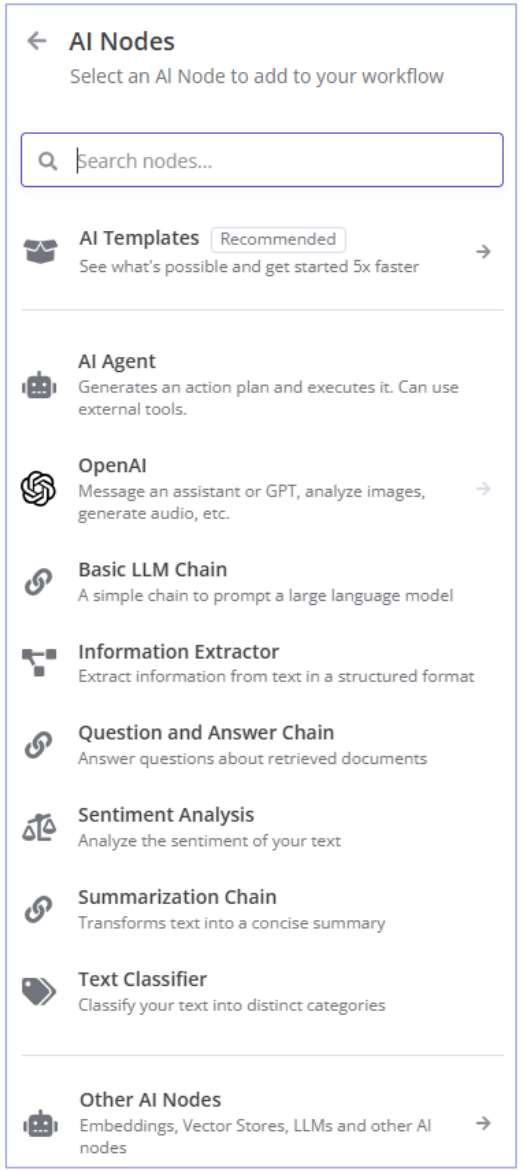
Using Chat


Chat
Session
How good is the RSA audience?
The RSA (Royal Society for the Encouragement of Arts, Manufactures and Commerce) is a UK-based organization that aims to promote social progress through ideas, innovation, and practical application. The quality of the RSA's events, including their audience, can vary depending on several factors such as the specific event, the topic, and the audience.
Here are some general observations about the RSA audience:
1. Innovators and changemakers: The RSA attracts a diverse audience of innovators, entrepreneurs, policymakers, and thought leaders who are passionate about creating positive social change.
2. Highly engaged: Attendees at RSA events are often highly engaged and motivated to learn from each other and from the speakers. This can lead to lively discussions, networking opportunities, and a sense of community among attendees.
3. Diverse perspectives: The RSA audience tends to be diverse in terms of age, background, occupation, and interests. This diversity brings different perspectives and insights to the table, enriching the events and discussions.
4. Crossover with other groups: RSA events often attract crossover audiences from related fields, such as tech, policy, arts, and social impact. This can lead to interesting connections and collaborations between attendees.
However, some potential drawbacks of the RSA audience include:
Type a message, or press 'up' arrow for previous one
Latest Logs from Basic LLM Chain node
Basic LLM Chain
Ollama Chat Model
Ollama Chat Model
15802ms | Started at 12:58:50 PM | ~394 Tokens
Input
Human: How good is the RSA audience?
Output
The RSA (Royal Society for the Encouragement of Arts, Manufactures and Commerce) is a UK-based organization that aims to promote social progress through ideas, innovation, and practical application. The quality of the RSA's events, including their audience, can vary depending on several factors such as the specific event, the topic, and the audience.
Here are some general observations about the RSA audience:
1. Innovators and changemakers: The RSA attracts a diverse audience of innovators, entrepreneurs, policymakers, and thought leaders who are passionate about creating positive social change.
2. Highly engaged: Attendees at RSA events are often highly engaged and motivated to learn from each other and from the speakers. This can lead to lively discussions, networking opportunities, and a sense of community among attendees.
3. Diverse perspectives: The RSA audience tends to be diverse in terms of age, background, occupation, and interests. This diversity brings different perspectives and insights to the table, enriching the events and discussions.
4. Crossover with other groups: RSA events often attract crossover audiences from related fields, such as tech, policy, arts, and social impact. This can lead to interesting connections and collaborations between attendees.
However, some potential drawbacks of the RSA audience include:

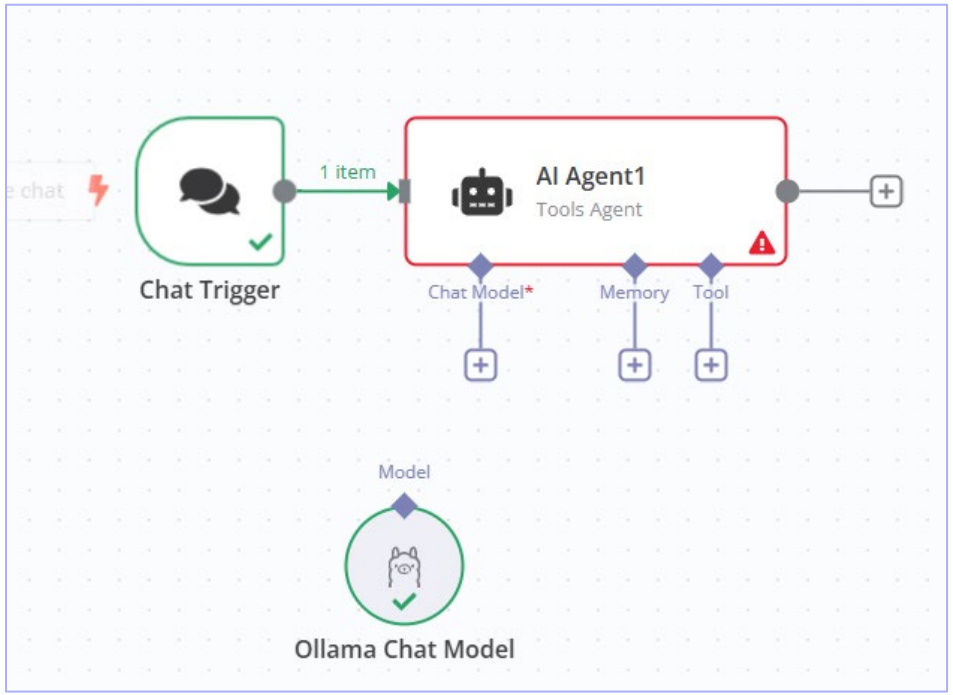
Additional AI/LLM Functionality

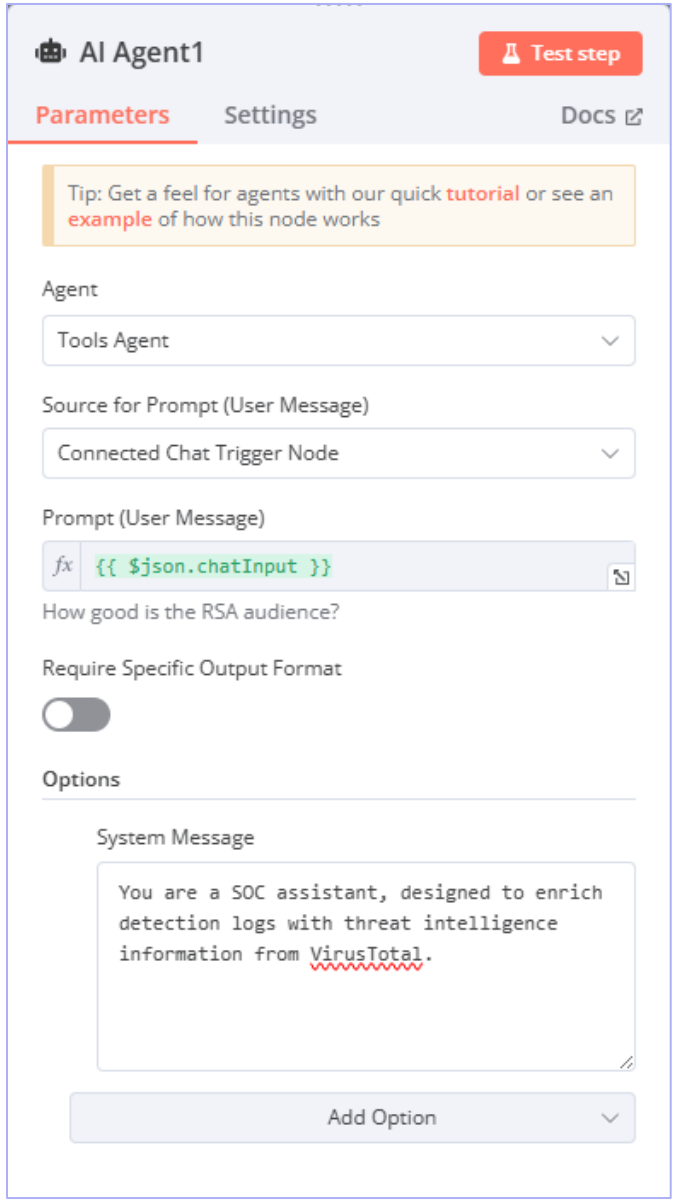
Common AI Nodes

- Agent
- Chat
- LLM chain
- Text classifier
- Embeddings
- Vector stores
- Much more!

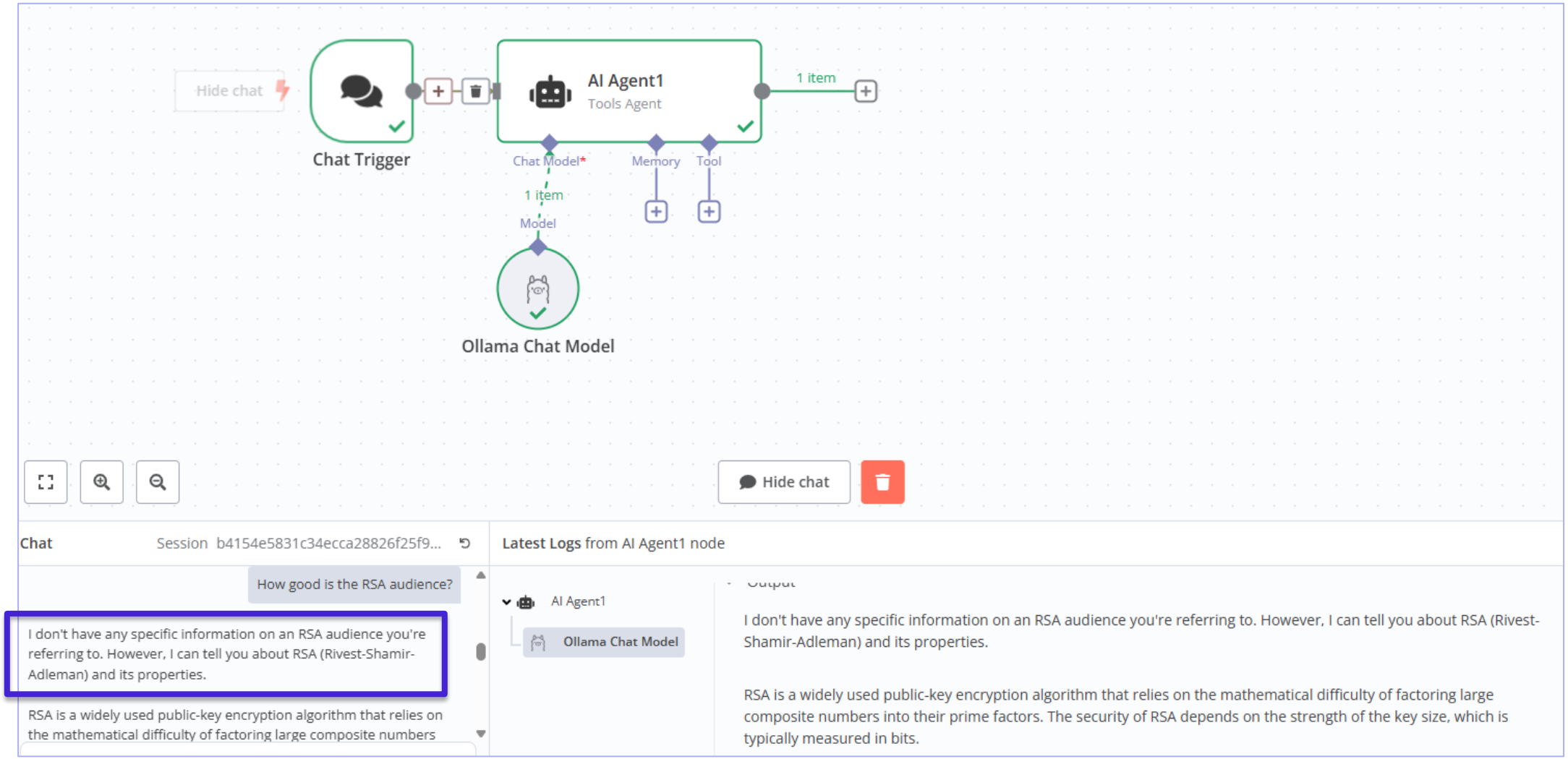

AI Nodes
Select an AI Node to add to your workflow
Search nodes...
AI Templates
Recommended
See what's possible and get started 5x faster
AI Agent
Generates an action plan and executes it. Can use external tools.
OpenAI
Message an assistant or GPT, analyze images, generate audio, etc.
Basic LLM Chain
A simple chain to prompt a large language model
Information Extractor
Extract information from text in a structured format
Question and Answer Chain
Answer questions about retrieved documents
Sentiment Analysis
Analyze the sentiment of your text
Summarization Chain
Transforms text into a concise summary
Text Classifier
Classify your text into distinct categories
Other AI Nodes
Embeddings, Vector Stores, LLMs and other AI nodes

Changing to AI Agent

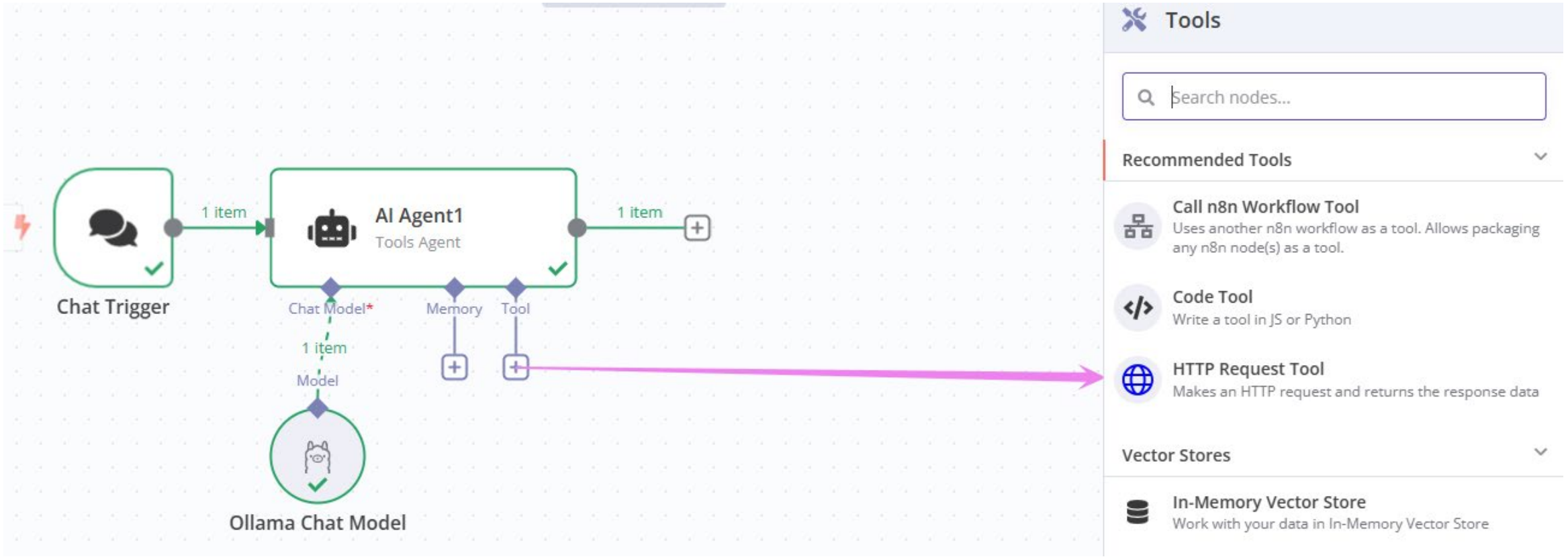



AI Agent1
Test step
Parameters
Settings
Docs
Tip: Get a feel for agents with our quick tutorial or see an example of how this node works
Agent
Tools Agent
Source for Prompt (User Message)
Connected Chat Trigger Node
Prompt (User Message)
{{ $json.chatInput }}
How good is the RSA audience?
Require Specific Output Format
Options
System Message
You are a SOC assistant, designed to enrich detection logs with threat intelligence information from VirusTotal.
Add Option

Updated Response with Prompt Instructions

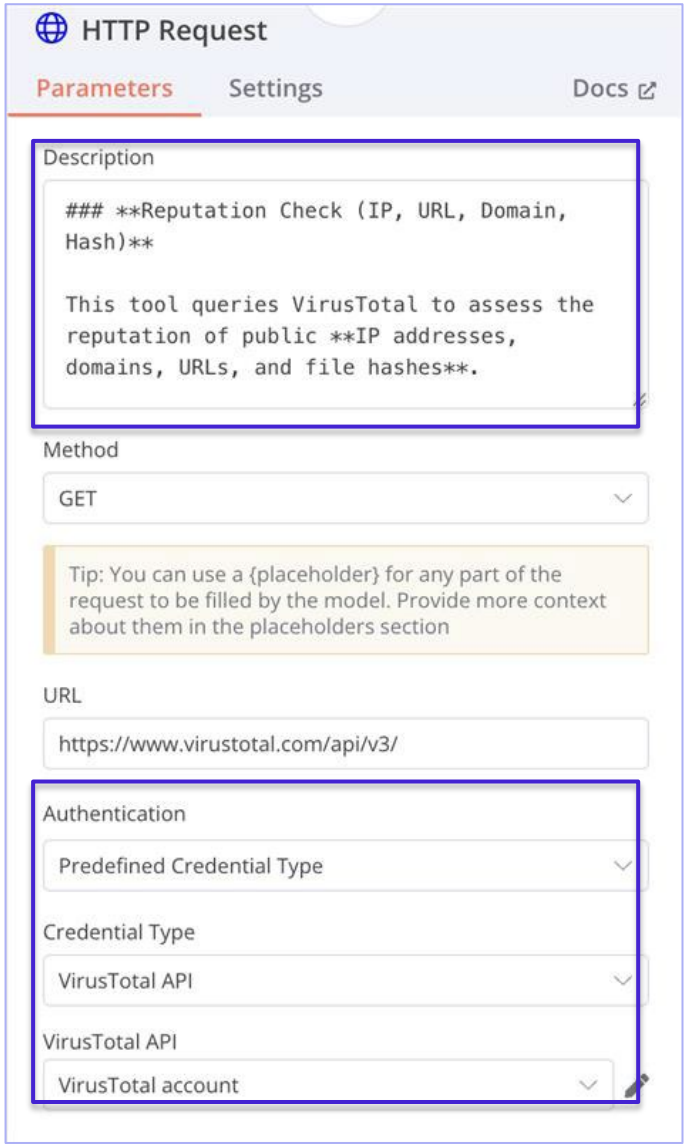

Hide chat
Chat Trigger
AI Agent1
Tools Agent
1 item
Chat Model*
Memory
Tool
1 item
Model
Ollama Chat Model
Hide chat
Chat
Session b4154e5831c34ecca28826f25f9...
How good is the RSA audience?
I don't have any specific information on an RSA audience you're referring to. However, I can tell you about RSA (Rivest-Shamir-Adleman) and its properties.
RSA is a widely used public-key encryption algorithm that relies on the mathematical difficulty of factoring large composite numbers
Latest Logs from AI Agent1 node
AI Agent1
Ollama Chat Model
I don't have any specific information on an RSA audience you're referring to. However, I can tell you about RSA (Rivest-Shamir-Adleman) and its properties.
RSA is a widely used public-key encryption algorithm that relies on the mathematical difficulty of factoring large composite numbers into their prime factors. The security of RSA depends on the strength of the key size, which is typically measured in bits.

Adding in the Threat Intelligence Tool



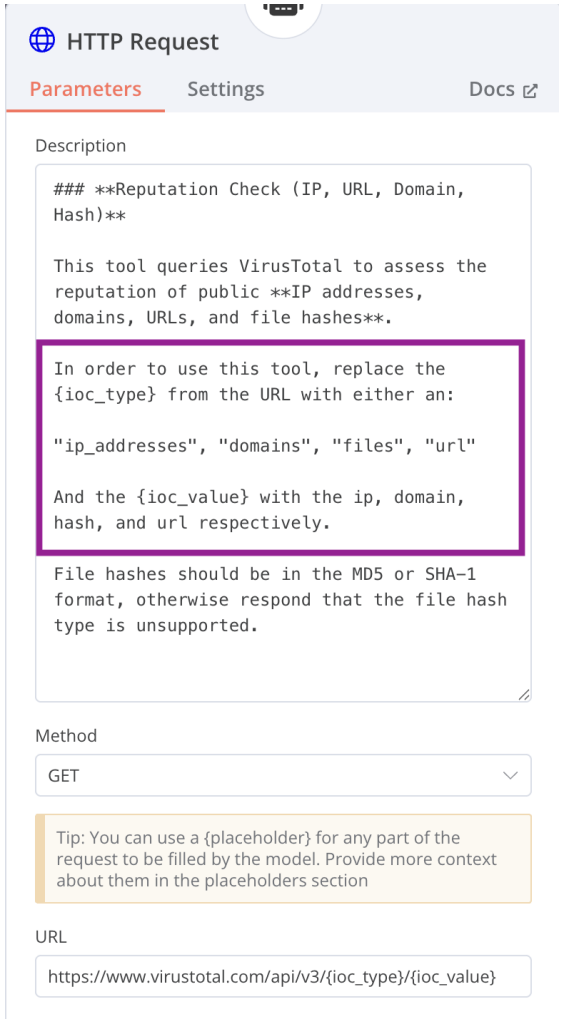
Predefined Auth (i.e. VirusTotal)

- Just enter your API key from VT!
- Add a description (semi-prompt)


HTTP Request
Parameters
Settings
Docs
Description
Reputation Check (IP, URL, Domain, Hash)
This tool queries VirusTotal to assess the reputation of public **IP addresses, domains, URLs, and file hashes**.
Method
GET
Tip: You can use a {placeholder} for any part of the request to be filled by the model. Provide more context about them in the placeholders section
URL
https://www.virustotal.com/api/v3/
Authentication
Predefined Credential Type
Credential Type
VirusTotal API
VirusTotal API
VirusTotal account

Edit Prompt to Format URL/API Request

- **The tool agent prompt can be highly specific:**
  - i.e. handling “variables”
  - {ioc_type} and {ioc_value}


HTTP Request
Parameters
Settings
Docs
Description
Reputation Check (IP, URL, Domain, Hash)
This tool queries VirusTotal to assess the reputation of public **IP addresses, domains, URLs, and file hashes**.
In order to use this tool, replace the {ioc_type} from the URL with either an:
"ip_addresses", "domains", "files", "url"
And the {ioc_value} with the ip, domain, hash, and url respectively.
File hashes should be in the MD5 or SHA-1 format, otherwise respond that the file hash type is unsupported.
Method
GET
Tip: You can use a {placeholder} for any part of the request to be filled by the model. Provide more context about them in the placeholders section
URL
https://www.virustotal.com/api/v3/{ioc_type}/{ioc_value}

Testing VirusTotal Tool

Our Tool

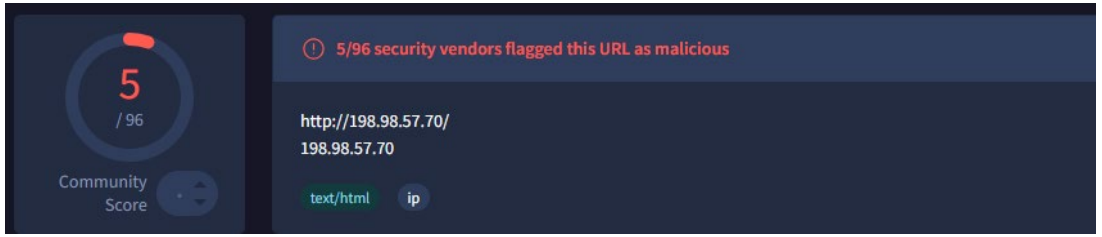
What is the reputation of IP Address 198.98.57.70

The provided JSON object appears to be a malware detection result from a reputable source, such as VirusTotal or Malwarebytes. Here's a breakdown of the key information:

1. **Malware Classification**:
   - `malicious` : 5 samples identified as malicious
   - `suspicious` : 1 sample identified as suspicious (not fully confirmed)
   - `undetected` : 30 samples detected but not fully analyzed
   - `harmless` : 58 samples identified as harmless
2. **Network Information**:
   - `network` : The IP range associated with the malware, which is `198.98.48.0/20`
3. **Organization Information**:
   - `NetRange` , `NetName` , and `NetHandle` provide information about the organization behind the network (PONYNET-06)
4. **Whois Data**:
   - The whois data provides detailed information about the organization, including their name, address, city, state, postal code, country, and registration dates.
5. **Detection Methods**:
   - `whois` checks if the IP is listed in a Whois database
   - `last_analysis_stats` displays the results of previous malware analysis

Based on this information, it appears that the malware was detected from an IP address associated with

VirusTotal



Enriching Logs with Threat Intelligence

- **End goal:**
  Use a vector store to store SIEM logs and enrich dynamically with threat intelligence
- **Final step:**
  Integrate vector store

Log Source Vector Store

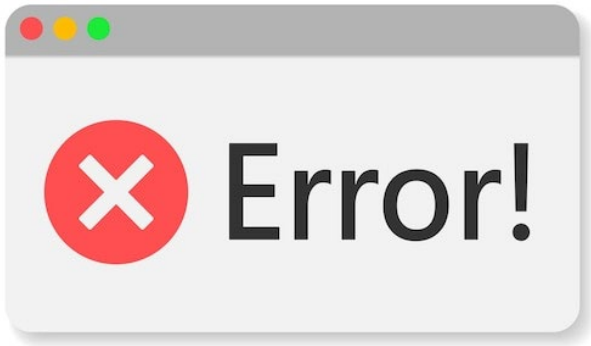
- **Add a new “manual” trigger to integrate vector store**
- **Add a new vector store** (i.e. Qdrant, comes with this build)
- **Add a Set Node** – data transformation node
- **Create a Qdrant collection with the sample log (not shown)**
  - Browse to localhost:<qdrant port>/dashboard
  - Quickstart
  - Create a collection
- **Load collection in our Set Node**
- **Add embeddings model for Ollama** (i.e. nomic-embed-text)
- **Add Default Data Loader** (document)
- **Add Text-Splitter to optimize embedding chunks**

Steps Cont'd

- Add a new Qdrant Vector Store Tool (for log retrieval)
- Add same embeddings model to Vector Store
- Update the AI agent prompt to incorporate knowledge of log source

Inevitable Debugging

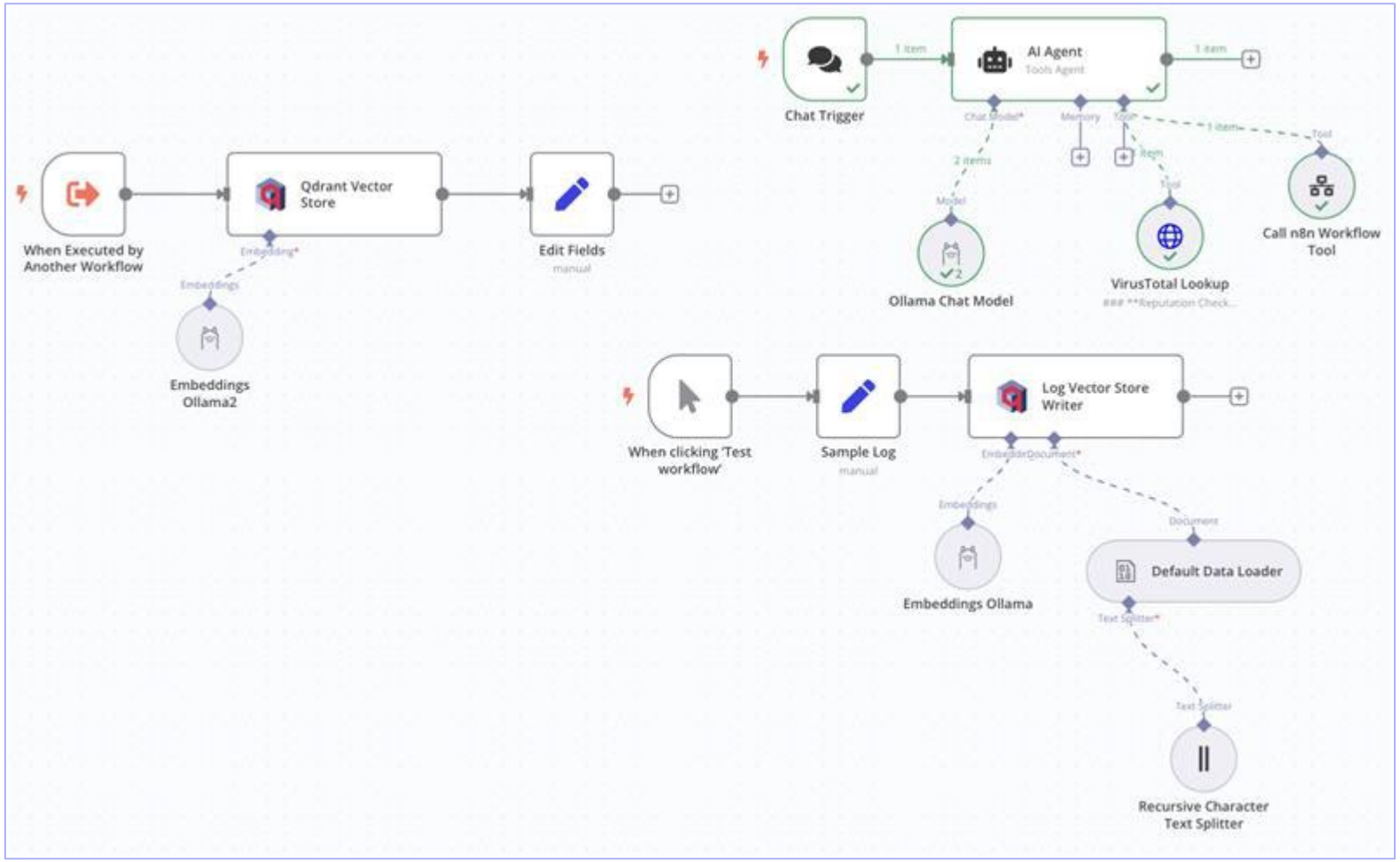
- “non string tool message content is not supported”
- Existing known issue with ollama string/json transformations
- Worked around this by calling another workflow and manually transforming the vector store output into a string



Our Sample Log

destinationServiceName=CrowdStrike dproc=FDR Items explorer </>
{"eid":118,"SensorId":"3f1f8807585a4c2092f3abcf54fc715b","Tactic":"Post-Exploit","CustomerIdString":"922f1fcc87cd4a6d9eb54d93c710f8ad","EventType":"Event_ExternalApiEvent","ParentProcessId":52984911733,"FileName":"gpupdate.exe","UTCTimestamp":1741816409000,"Malicious artifacts were identified in memory. The process may be hijacked or malicious. Adversaries execute malicious code in memory to evade detection. Investigate the process and surrounding events.","ParentCommandLine":"\"Code.exe\"","timestamp":"2025-03-12T21:53:29Z","FilePath":"\\Device\\HarddiskVolume2\\Windows\\System32","UserName":"Administrator","MD5String":"3e8b1f922818935d3de355fc6ab500b2","SeverityName":"High","PatternDispositionDescription":"Prevention, process killed.","Severity":4,"DetectId":"ldt:3f1f8807585a4c2092f3abcf54fc715b:17180241268","PatternDispositionValue":16,"ExternalApiType":"Event_DetectionSummaryEvent",**"SHA256String:":"226a723ffb4a91d9950a8b266167c5b354ab0db1dc225578494917fe53867ef2"**"NetworkAccesses":[{"AccessTimestamp":1741805109**"RemoteAddress:":"104.21.50.73","LocalAddress":"10.2.0.166"**calPort":55308,"RemotePort":443,"IsIPV6":false,"Protocol":"TCP","AccessType":0}],"Nonce":1,"Objective":"Falcon Detection Method","CommandLine":"C:\\Windows\\system32\\gpupdate.exe","MACAddress":"00-50-56-95-94-49","GrandparentCommandLine":"\"C:\\Users\\Administrator\\Downloads\\NotE\\VSCode-win32-x64-1.95.1\\VSCode-win32-x64-1.95.1\\Code.exe\"","ProcessStartTime":1741805109,"ComputerName":"USR670S3207","DetectName":"Suspicious Activity","QuarantineFiles":[{"ImageFileName":"\\Device\\HarddiskVolume2\\Users\\Administrator\\Downloads\\NotE\\VSCode-win32-x64-1.95.1\\VSCode-win32-x64-1.95.1\\CRYPTBASE.DLL","SHA256HashData":226a723ffb4a91d9950a8b266167c5b354ab0db1dc225578494917fe53867ef2"}],"AgentIdString":"3f1f8807585a4c2092f3abcf54fc715b","ProcessId":53043405030,"Technique":"Malicious Tool Execution","cid":"922f1fcc87cd4a6d9eb54d93c710f8ad"}

The Final Model



RSAC™ 2025 Conference

Lessons Learned
Triumphs and Challenges

Many Voices.
One Community.

Success Stories and Challenges

Successes

- Increased automation and efficiency
- Faster threat detection and response
- Improved analyst productivity

Challenges

- Keeping LLM responses relevant and factual
- Managing hallucinations and inaccuracies
- Balancing automation with human oversight
- Getting the prompt right and using a large enough model

Take Action This Week

- **Pick one** use case and apply LLM automation
- **Experiment** with prompt engineering in a security context
- **Ensure** proper security controls are in place before deploying AI in workflows
- **Publish** your workflow/template!

Resources

- **n8n starter kit** (https://github.com/n8n-io/self-hosted-ai-starter-kit)
- **Langchain tutorial** (https://academy.langchain.com/courses/take/intro-to-langgraph/texts/58238105-getting-set-up)
- **Intro to LLMs** (https://www.youtube.com/watch?v=zjkBMFhNj_g)
- **n8n templates / workflows** (https://n8n.io/workflows/)

RSAC™ 2025 Conference

Many Voices.
One Community.

Thank You

For more information: